The Ax-Grothendieck Theorem

Main Reference:

[1]. Jean-Pierre Serre. How to use finite fields for problems concerning infinite fields.

[2]. Terrace Tao. Infinite fields, finite fields, and the Ax-Grothendieck theorem. (personal blog)

[3]. Anand Pillay. Model theory (short article on Notices of AMS)

Sometimes algebraic statements over fields of char. zero, such as $\mathbb{C}$, can be deduced from their positive char. counterparts such as $\mathbb{F}_p$.

- "fields of char >0 can partially model fields of char 0."
- If a certain algebraic statement failed over (say) $\mathbb{C}$, then there should be a "finitary algebraic obstruction" that "witnesses" this failure over $\mathbb{C}$.

Thm: (Ax-Grothendieck, 1960s) Let X be an alg. variety over an alg. closed field k. If a morphism $f: X \to X$ is inj., then it is bij.

Let's look at a special case first:

Thm': If $P: \mathbb{C}^n \to \mathbb{C}^n$ is an inj. polynomial map, then P is bij.

Proof via finite fields

Observations:

- 1) The theorem is trivial in finite field case.
- 2) Pass from a finite field F to its algebraic closure $\bar{F}$:

Prop': If $\underline{P}: \bar{F}^n \to \bar{F}^n$ is an inj. polynomial map, then $\underline{P}$ is bij. (strong version)

Pf: ① By Hilbert's Nullstellensatz. Suppose $\underline{P}: \bar{F}^n \to \bar{F}^n$ is inj. but not surj. Then from inj. of P we see the alg. system

$$\underline{P}(\underline{x}) = \underline{P}(\underline{y}) \; ; \; \underline{x} \neq \underline{y}$$

has no soln. over $\bar{F}$. $\Rightarrow \forall j = 1, \cdots, n$, $\exists$ an alg. identity of the form

$$(\underline{P}(\underline{x}) - \underline{P}(\underline{y})) \cdot \underline{Q}_j(\underline{x}, \underline{y}) = (x_j - y_j)^{r_j} \qquad (1)$$

for some $r_j \geq 1$, some poly. $\underline{Q}_j: \bar{F}^n \times \bar{F}^n \to \bar{F}^n$.

Similarly, lack of surj. means $\exists \, \underline{z}_0 \in \bar{F}^n$, s.t. the alg. system

$$\underline{P}(\underline{x}) = \underline{z}_0$$

has no soln. over $\bar{F}$. $\Rightarrow \exists$ an alg. identity of the form

$$(\underline{P}(\underline{x}) - \underline{z}_0) \cdot \underline{R}(\underline{x}) = 1 \qquad (2)$$

for some poly. $\underline{R}: F^n \to F^n$.

Fix $\underline{Q}_j, \underline{R}, \underline{z}_0$ as above, and let E be the subfield of $\bar{F}$ generated by F & all coefficients of $\underline{P}, \underline{Q}_j, \underline{R}, \underline{z}_0$, then E/F is finite, and our counterexample $\underline{P}$ descends from $\bar{F}$ to $\underline{R}$ $\lightning$.

② A much quicker way: $\forall \underline{b} \in \bar{F}^n$. Let E be the subfield of $\bar{F}$ generated by F & all coefficients of $\underline{P}$ and $\underline{b}$, then $\underline{P}: E^n \to E^n$ is bij., so $\exists \underline{a} \in E^n \subset \bar{F}^n$. s.t. $\underline{P}(\underline{a}) = \underline{b}$. □.

- 3) Moving from fields of char. >0 to fields of char. 0: algebraic & model-theoretic approaches.

A. Algebraic approach

Pf of Thm 1': Since $\mathbb{C}$ is alg. closed, we may invoke the Nullstellensatz as before and find witnesses (1) & (2) for some $\underline{Q}_j, \underline{z}_0, \underline{R}$.

Let $\mathbb{Q}(C) (= E)$ be the subfield of $\mathbb{C}$ gen. by $\mathbb{Q}$ & all the coefficients C of $\underline{P}, \underline{Q}_j, \underline{z}_0, \underline{R}$. There are several ways to descend E to a finite field, e.g.

* ① (Serre) quotient the ring $\mathbb{Z}[C]$ by a max'l ideal;

② using a general mapping theorem of Van H. Vu, Melanie N. Wood & Philip M. Wood. Sketch:

E has a finite transcendence basis $\alpha_1, \cdots, \alpha_m$ over $\mathbb{Q}$. By the primitive el't thm, E can be expressed as $\mathbb{Q}(\alpha_1, \cdots, \alpha_m, \beta)$ for some β alg. over $\mathbb{Q}(\alpha_1, \cdots, \alpha_m)$. I.p., all the coefficients C are rational combinations of $\alpha_1, \cdots, \alpha_m, \beta$. Now,

- By rationalizing $\rightsquigarrow$ can ensure that the denominators of the expressions of these coefficients are in the ring $\mathbb{Z}[\alpha_1, \cdots, \alpha_m]$
- Dividing β by some power of the product of these denomi-

nators $\rightsquigarrow$ all coefficients in C lie in $\mathbb{Z}[\alpha_1, \cdots, \alpha_m, \beta]$

- $\mathbb{Z}[\alpha_1, \cdots, \alpha_n, \beta] \cong \mathbb{Z}[a_1, \cdots, a_n][b] / f(b)$, f being the minimal polynomial of β

Hence, (1) & (2) transfer to this ring. Pick a large prime p, and map $a_1, \cdots, a_m$ to random elts of $\mathbb{F}_p$, the image $\bar{f}$ of f is often non-degenerate, so we can then map b to a root of this image in a finite extension of $\mathbb{F}_p$. $\Rightarrow$ This descends (1) & (2) to a finite field, as desired. $\square$.

Proof via model theory

- A detour to ultra products -

Let X be a set, a filter $\mathcal{F}$ on X is a set of subsets of X with the following conditions:

(F_1) $\emptyset \notin \mathcal{F}$, $\mathcal{F} \neq \emptyset$;

(F_2) If $A, B \in \mathcal{F}$, then so is $A \cap B$;

(F_3) If $A \in \mathcal{F}$, $B \supset A$, then $B \in \mathcal{F}$.

E.g. • X top. space, $x \in X$, $\mathcal{F} = \{$ nbhds of x in $X\}$

• $X = \mathbb{N}$, $\mathcal{F} = \{A \subset X$ cofinite$\}$

A filter $\mathcal{F}$ on X is an ultrafilter if $\forall A \subset X$, either A or its complement is in $\mathcal{F}$. If $\exists x \in X$, s.t. $\mathcal{F} = \{A : x \in A\}$, then $\mathcal{F}$ is called a principal ultrafilter.

Rmk: (1) A filter $\mathcal{F}$ is a principal ultrafilter iff contains a singleton; an ultrafilter is principal iff it contains a finite set.

(2) Originally, filters were invented by H. Cartan to define the most general form of limit in topology: given a filter $\mathcal{F}$ on a top. space X, an elt $x \in X$ is a limit of $\mathcal{F}$ iff $\mathcal{F} \supset$ {nbhds of x}.

E.g., $X = \mathbb{R}$, $(x_n)_{n\in\mathbb{N}}$ seq. in $\mathbb{R}$, $\mathcal{F} = \{A \subset \mathbb{R} : A$ contains almost all x_n's (except for finitely many)$\}$. Then, $\mathcal{F}$ is a filter on $\mathbb{R}$ & $\mathcal{F}$ has limit $x \in \mathbb{R} \Leftrightarrow \lim_{n\to\infty} x_n = x$.

(3) I.p., one can show: X is cpt $\Leftrightarrow$ every ultrafilter on X converges (and X is Hausdorff).

Now, let $\mathcal{L}$ be a language, given a set I and let $M_i, i \in I$ be a family of $\mathcal{L}$-structures with $M_i \neq \emptyset$.

(If you're not familiar with model theory, just think of $\mathcal{L}$ in our case to be formal things you need to give X a gp./ring/ordered set/$\cdots$ str., e.g. $\mathcal{L}_{ring} = \{1, 0, +, -, \times\}$, and $\mathcal{L}$-str. is a set which has such things.)

Let $\mathcal{F}$ be an ultrafilter on I, then the ultraproduct $\prod_{\mathcal{F}} M_i$ is the $\mathcal{L}$-str. defined as follows:

(1) $M = \prod_{\mathcal{F}} M_i = \prod_{i\in I} M_i / \sim$, where

$$(m_i)_{i\in I} \sim (n_i)_{i\in I} \Leftrightarrow \{i \in I : m_i = n_i\} \in \mathcal{F}.$$

(2) If c is a constant symbol, then

$$c_M = [(c_{M_i})_{i\in I}],$$

where $[-]$ represents the class of $-$ in $\prod_{i\in I} M_i$ modulo $\sim$.

(3) If (r, n) (resp. (f, n)) is a relation (resp. function) symbol, then

$$r_M = \{ \text{class of } (x_1, \dots, x_n) \in (\prod_{i \in I} M_i)^n \text{ s.t. } \{ i \in I : (x_{1,i}, \dots, x_{n,i}) \in r_{M_i} \} \in \mathcal{F} \}$$

(resp.

$$f_M([x_1], \dots, [x_n]) = [(f(x_i))_{i \in I}].)$$

Intuitively, $\mathcal{F}$ corresponds to sets containing "almost all" $i \in I$, so a binary rel'n $\leq$ will hold between $x = (x_i)$ and $y = (y_i)$ iff "$x_i \leq y_i$ for almost all i".

In our case, $\mathcal{L} = \mathcal{L}_{ring}$, I will be the set P of all primes, and for each $p \in P$, M_p will be the alg. closure $\bar{\mathbb{F}}_p$ of $\mathbb{F}_p$.

For another (not so interesting) example, if $\mathcal{F}$ is principal, i.e., $\mathcal{F} = \{ A \subseteq I : i_0 \in A \}$, then $\prod_{i \in I} M_i$ realizes M_{i_0}.

Here is the key property we need:

Thm (Łoś): $\mathcal{L}$, I, $(M_i)_{i \in I}$, and $\mathcal{F}$ as above. For any $\mathcal{L}$-formula $\varphi(x_1, \dots, x_n)$ and $a \in (\prod_{\mathcal{F}} M_i)^n$, we have $a \in \varphi(\prod_{\mathcal{F}} M_i)$ iff $\{ i \in I : M_i \models \varphi \} \in \mathcal{F}$.

I.p., if T is an $\mathcal{L}$-theory and $M_i \in T$, $\forall i$, then we have $\prod_{\mathcal{F}} M_i \models \varphi$.

In our case, there is a theory T_{ACF} of algebraic closed fields, so it follows that $\prod_{\mathcal{F}} \bar{\mathbb{F}}_p$ is automatically an ACF. More interestingly, C is of characteristic 0 if $\mathcal{F}$ is not principal. (To see this, for each $p \in P$, a field F is of char. p iff $F \models \varphi_p$, where φ_p is the formula: $\forall x \; \underbrace{x \cdot x \cdot \dots \cdot x}_{p \text{ times}} = 1$. But φ_p holds only for $M_p = \bar{\mathbb{F}}_p$.)

$\sim =: C$

We are finally ready to prove Thm'.

Proof of Thm' (II): Recall we have shown that Thm' holds for $X = \overline{\mathbb{F}}_p^n$ (Prop'), by Łoś theorem above we see Thm' also holds for our ultraproduct $C = \prod_{\mathcal{F}} \overline{\mathbb{F}}_p$ (Injectivity implies surjectivity can be expressed as a first-order formula: Write a finite set of polynomials in n variables & degree $\leq d$ in multi-index notation as

$$f_i(\underline{x}) = \sum_{\alpha} a_\alpha^i \underline{x}^\alpha,$$

then we write

$$\varphi_{n,d} := \text{“} \forall a_\alpha^i \left(\forall \underline{x}, \underline{y} \left(\bigwedge_{i=1}^{n} \sum_\alpha a_\alpha^i \underline{x}^\alpha = \sum_\alpha a_\alpha^i \underline{y}^\alpha \rightarrow \underline{x} = \underline{y} \right) \rightarrow \left(\forall \underline{y}, \exists \underline{x} \bigwedge_{i=1}^{n} \sum_\alpha a_\alpha^i \underline{x}^\alpha = \underline{y} \right) \right) \text{”}$$

For all n, d, $\varphi_{n,d}$ holds in finite fields & $\overline{\mathbb{F}}_p$, and hence in C.)

So, it suffices to show that $C \cong \mathbb{C}$, for which we borrow the following fact —

Fact: Alg. closed fields of char. 0 are Iso. Iff they have the same cardinality.

On the one hand,

$$\left| \prod_{\mathcal{F}} \overline{\mathbb{F}}_p \right| \leq \aleph_0^{\aleph_0} = 2^{\aleph_0} = |\mathbb{C}|$$

where "$=$" is because we can construct an injective map f from $\mathbb{N}^{\mathbb{N}}$ to $\{0,1\}^{\mathbb{N}}$ as follows: let $a = (a_n)_{n \in \mathbb{N}}$ be a seq. of nat. numbers, then we define

$$f(a) = (\underbrace{1, \cdots, 1}_{a_0 \text{ copies}}, 0, \underbrace{\cdots, 0}_{a_1 \text{ copies}}, \underbrace{1, \cdots, 1}_{a_2 \text{ copies}}, \cdots).$$

On the other hand, we construct an injective map

$$\mathbb{C} \xrightarrow{\sim} \mathbb{R} \to \prod_P \bar{\mathbb{F}}_p \twoheadrightarrow \hat{\prod}_{\mathcal{F}} \mathbb{F}_p = C$$

(equal card.)

by the following recipe: (i) There exists a family $(f_t)_{t\in\mathbb{R}}$ of maps $f_t : P \xrightarrow{\sim} \mathbb{N} \to \mathbb{Q}$ s.t. $\forall\, t \neq s \in \mathbb{R}$, the set

$$\{p \in P : f_t(p) = f_s(p)\}$$

is finite (for instance, take for each p a seq. of rational numbers converging to t). (ii) Note that $\bar{\mathbb{F}}_p = \bigcup_{n\geq 1} \mathbb{F}_{p^n}$, we see that $|\bar{\mathbb{F}}_p| = |\mathbb{Q}|$. So let $\nu_p : \mathbb{Q} \to \bar{\mathbb{F}}_p$ be a bijection, we have then constructed a family of maps $g_t : P \to \bar{\mathbb{F}}_p$, $p \mapsto \nu_p(f_t(p))$, and hence a map

$$\begin{aligned} \mathbb{R} &\to \prod_P \bar{\mathbb{F}}_p \twoheadrightarrow \hat{\prod}_{\mathcal{F}} \mathbb{F}_p \\ t &\mapsto (g_t(p))_p \mapsto [(g_t(p))_p] \end{aligned}$$

It is inj. by construction. (Indeed, by Rmk (i) $(a_i) = (b_i)$ in C only if a_i & b_i agree on an infinite set, but the set

$$\{p \in P : g_t(p) = g_s(p)\}$$

is finite.) $\square$.

Proof via complex analysis (Rudin's proof)

Now we give Rudin's proof[4], which relies on some Galois theory and the top. str. of $\mathbb{C}$.

Lem[3]: Let $\Omega \overset{\text{open}}{\subset} \mathbb{C}^n$, $f : \Omega \to \mathbb{C}^n$ inj., holo. Then, the Jacobian of f is non-degenerate. i.e., $\det Df(\underline{z}) \neq 0$. $\forall\, \underline{z} \in \Omega$.

Pf (argument of Rosay[5]): By induction, the case $n=1$ is clear (See for instance Prop[VIII] 1.1, P. 206-207 of Stein's book.[6])

Suppose now the claim is proved for $n-1$, and suppose by contradiction that $\det Df(\underline{z}_0) = 0$ for some $\underline{z}_0 \in \Omega$. Claim:

$$Df(\underline{z}_0) = \underline{0}$$

[Indeed, if $\exists\, 1 \le i, j \le n$, s.t. $\frac{\partial}{\partial z_j} f_i(\underline{z}_0) \neq 0$, then WLOG we may assume $i = j = 1$ and normalize $z_0 = f(z_0) = 0$. Then, the map $h: \underline{z} \mapsto (f_1(\underline{z}), z_2, \cdots, z_n)$ is holo. with non-degenerate Jacobian at 0, and is thus loc. invertible at 0. The map $f \circ h^{-1}$ is then holo. at 0 & preserves the z_1 coord., and therefore descends to an inj. holo. map on a nbhd of the origin of $\mathbb{C}^n$, so its Jacobian is non-degenerate by induction hypothesis $\lightning$.]

We have shown that Df vanishes on the zero set

$$\{\det Df = 0\},$$ if it is non-empty

which is an analytic variety of codim 1 (In our case, f is a polynomial, so it is an alg. variety). Thus, f is loc. const. on this variety, which contradicts inj. $\square$.

Cor: Inj. holo. maps from $\mathbb{C}$ to $\mathbb{C}$ are open.

Proof of Thm 1' (III): Let $\underline{P}: \mathbb{C}^n \to \mathbb{C}^n$ be a polynomial, and let K be the field generated by $\mathbb{Q}$ and the coefficients of $\underline{P}$. Denote by $K(\underline{z})$ $K(z_1, \cdots, z_n)$ the extension of K by n indeterminates $z_1, \cdots, z_n$, and consider the subfield $K(\underline{P}(\underline{z})) \subset K(\underline{z})$.

Claim: $K(\underline{P}(\underline{z})) = K(\underline{z})$. [If not, then $\exists$ a non-trivial automorphism $\varphi: K(\underline{z}) \to K(\underline{z})$ that fixes $K(\underline{P}(\underline{z}))$. I.p., $\exists$ a

non-trivial rational (over K) combination $Q(z)/R(z)$ of z s.t. $P(Q(z)/R(z)) = P(z)$. ① Now, map z to a random complex number in $\mathbb{C}$, which will almost surely be transcendental over $K \Rightarrow P$ can not be inj. ↯

Since $K(P(z)) = K(z)$, $\exists$ rational fcns $Q_j(z)/R_j(z)$, $j=1,\cdots,n$, s.t. $z_j = Q_j(P(z))/R_j(P(z))$. We may assume that Q_j, R_j have no common factors. These give us poly. identities

$$Q_j(P(z)) = z_j R_j(P(z)), \ \forall j.$$

I.p., on the open set $P(\mathbb{C}^n) \subset \mathbb{C}^n$, the zero set of R_j is contained in that of Q_j, which is impossible. ② So $R_j \neq 0$ on $P(\mathbb{C}^n)$. Thus, the polynomials

$$R_j \circ P$$

(why?) has no zeros $\Rightarrow$ constant. We may then normalize so that $R_j \cdot P = 1$. Thus, we now have $z = Q(P(z))$ for some poly Q, which implies that $w = P(Q(w)), \forall w \in P(\mathbb{C}^n)$. But both w and $P(Q(w))$ are polynomials, they must agree on all of $\mathbb{C}^n$, so P is bij. □.

Rmk: (1) Cor 4 is also a special case of the well-known invariance of domain theorem.

(2) Rudin's proof shows also that the inverse of P is again a polynomial. This may be regarded as a small step toward the Jacobian conjecture: If $P: \mathbb{C}^n \to \mathbb{C}^n$ is an inj. poly. map whose Jacobian is a non-zero const., then P is a poly. auto-

morphism of $\mathbb{C}^n$.

①: Write $P_j(\underline{z}) = w_j$, then it suffices to show:

$$K(w_1, \cdots, w_n) = K(z_1, \cdots, z_n, w_1, \cdots, w_n).$$

If not, $\exists \varphi \in Gal^F(E/F)$, $\varphi \neq Id$. Let $\underline{x} = \varphi(\underline{z})$, then

$$\underline{P}(\underline{z}) = \varphi(\underline{P}(\underline{z})) = \underline{P}(\varphi(\underline{z})) = \underline{P}(\underline{x}),$$

where $\underline{x} \neq \underline{z}$.

②: K is countable, so there are only countably many polynomials with coefficients in K $\Rightarrow$ The union of their zero-sets is thus a countable union of closed sets without interior, hence can not cover $\mathbb{C}^n$.

Finally, we remark that both proofs (I) & (II) of Thm' can be generalized without much difficulty to obtain Thm'.

[4]. Elias M. Stein, and Rami Shakarchi. Complex analysis. Vol. 2. Princeton University Press, 2010.

[5]. Jean-Pierre Rosay. "Injective holomorphic mappings." The AMS Monthly 89.8 (1982): 587-588.

[6]. Walter Rudin. "Injective polynomial maps are automorphisms." The AMS Monthly 102.6 (1995): 540-543.

The (weak) Hilbert's Nullstallensatz states that for any prime ideal

$$\mathfrak{p} \subset F[x_1, \cdots, x_m],$$

where $m \geq 1$, $F = \bar{F}$ is alg. closed, there exists an elt $\underline{x} = (x_1, \cdots, x_m) \in F^m$ s.t. $f(\underline{x}) = 0$, $\forall f \in \mathfrak{p}$. We will present a model theoretic proof of this result here, the key ingredient is:

Thm: Let $\mathcal{L} = \{+, \cdot, -, 0, 1\}$ be the language of rings, and TACF be the theory of alg. cl. fields. Then TACF has quantifier elimination (i.e., every $\mathcal{L}$-formula is equivalent to a quantifier free one.

To illustrate this, let F be a model of TACF (so F has a str. of ACF), then the formula

$$\varphi(a, b, c) : \exists x,\ ax^2 + bx + c = 0$$

can be reduced to

$$\psi : (a \neq 0) \vee (b \neq 0) \vee (c = 0).$$

In general, a $\mathcal{L}$-formula with some free variables is an expression obtained by finitely many uses of the following rules

(i) for terms $t_1(\underline{v})$ & $t_2(\underline{v})$ ← In our case, polynomials

$$t_1(\underline{v}) = t_2(\underline{v})$$ w/ coefficients in F

is a formula;

(ii) connectors: $\wedge$ (and); $\vee$ (or); $\neg$ (negation); $\rightarrow$ (implies)

(iii) quantifiers: $\exists$, $\forall$

So, the Thm above shows that any formula can be reduced to $(F_1 = 0) \vee \cdots \vee (F_m = 0) \vee (F_{m+1} \neq 0) \cdots \vee (F_n \neq 0)$.

Cor: TACF is model-complete. (This holds for any theory w/ quantifier elimination) Namely, if $F_1 \subset F_2$ are alg. cl., then any "statement" that holds for F_1 will hold for F_2 and vice versa. (≈ we say sentence in model theory)

Cor: $TACF_p$ is complete. Any "statement" holds either for all alg. cl. fields or for no ---. (ACF of char. p)

Pf: Let F be an ACF of char p, then $F \supset \bar{\mathbb{F}}_p$, so a statement holds for $F \Leftrightarrow$ it holds for $\bar{\mathbb{F}}_p$, and the result follows. □.

Rmk: This is not true for the theory of fields. e.g., the statement "$\exists x\ (x^2+1=0)$" is valid in some fields but not all.

Now, we are ready to prove the weak Nullstellensatz:

By Hilbert's Basis Theorem, $\mathfrak{p} = (f_1, \cdots, f_R)$. Consider $K = \mathrm{Frac}(F[x_1, \cdots, x_m]/\mathfrak{p})$ and let $\bar{K}$ be its alg. closure. Clearly $F \subset K$. So

$$\text{"}\exists \underline{x}\ (f_1(\underline{x})=0) \wedge (f_2(\underline{x})=0) \cdots \wedge (f_R(\underline{x})=0)\text{"}$$

(view f_i's as formulas)

holds in $F \Leftrightarrow$ in K. But in K this is clear. we can just take $\underline{x} = (\bar{x}_1, \cdots, \bar{x}_m)$, where $\bar{x}_i$ is the image of x_i under

$$F[x_1, \cdots, x_m] \to F[x_1, \cdots, x_m]/\mathfrak{p} \hookrightarrow \bar{K}. \qquad \square.$$